DOCUMENT RESUME

ED 186 198 EC 012 031

AUTHOR McConnell, Beverly
TITLE Does Bilingual Education Work?
PUB DATE 3 Feb 79
NOTE 41p.: Paper presented at the Annual Meeting of the Southwest Educational Research Association (2nd, Houston, TX, February 3, 1979).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Bilingual Education: Comparative Testing: *Control Groups: Curriculum: Educational Objectives: English: Ethnicity: *Evaluation Methods: Hispanic Americans: Language Arts: *Language Dominance: Mathematics: *Program Effectiveness: Program Evaluation: Spanish Speaking: Test Results
IDENTIFIERS *American Institute for Research: *Individualized Bilingual Instruction Program

ABSTRACT

Two major evaluations of bilingual education differed significantly in evaluation design and results. The American Institute for Research (AIR) conducted the national impact study for bilingual education and gathered data from many programs. AIR based its study on comparisons of mathematics and language arts achievement of students in bilingual classrooms (with 74% non-English speaking or bilingual children) and control classrooms (with just 17% bilingual or non-English speaking children). The two groups were matched ethnically but not by language dominance. The design of the evaluation of the Individualized Bilingual Instruction (IBI) Program was based on evidence of how the subjects would have performed without the benefit of the program. These data, built up over years, were accumulated from actual pretests of students who enrolled in the program at different ages, thus providing an automatic match for ethnicity, socio-economic level, age, and language dominance. In general, IBI results indicated significantly higher performance for students in its bilingual program than did AIR for Title VII students as a whole. Appendices include the IBI data base from which the paper's findings were taken and technical information regarding IBI testing, data collection, and analysis procedures. (SE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Beverly B. McConnell
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED186198

RC 012031

DOES BILINGUAL EDUCATION WORK?

Beverly McConnell
Evaluator for Individualized Bilingual Instruction
(A Texas and Washington based bilingual
program administered through
Sunnyside School District 201,
Washington)

Presented Feb. 3, 1979 at the second annual meeting of the Southwest Educational Research Association, Houston, Texas

Those "present at the creation" when the bilingual education program was established by Congress in 1968 would probably have responded: "We don't know, but it is certainly worth a try." The "negative case" for bilingual education was demonstrated in the statistics that showed from 50% to 100% of language minorities dropped out of school prior to the 12th grade in a school system using only English as the medium of instruction.

Jose Cardenas, an early supporter of bilingual programs, described bilingual education as an "untested alternative," noting:

> Perhaps Hispanic minorities are so overwhelmingly in favor of bilingual education regardless of lack of evidence of its success because the experiences with past programs have been so negative . . . . If . . . the drop-out rate of Mexican American children in a South Texas school system is 90 percent, the parents cannot be blamed for strongly recommending an untested alternative.[1]

The 1970's, however, have become the "age of accountability" in education, and pressure has mounted on both federal officials and local programs to "come up with the answers" on the value of bilingual education.

Two types of efforts have been mounted by the U. S. Office of Education in this search for answers. One has been directed toward individual programs. Evaluation data from such programs have been screened by the Joint Dissemination Review Panel (JDRP) of USOE which attempts to determine that the results are valid and

reliable, as well as educationally significant. Such programs are "validated" and special attention given to the data they have produced. Individualized Bilingual Instruction (IBI), the program reported in this document, is one bilingual program validated by JDRP.[2]

The other type of evaluative effort has been directed toward classes of programs. These are nationwide "impact studies" examining a sampling of programs from a common funding source to determine if, collectively, they are more effective than diverse "other" educational approaches in meeting the needs of the target population. For bilingual education the national "impact study" was completed by American Institute for Research (AIR) with its final report issued in 1978.

The answer to the question: "Does Bilingual Education Work?" that emerges from the IBI data, and from the AIR Impact Study, is in sharp contrast as summarized below:

CONCLUSIONS FROM AIR AND FROM IBI DATA ON EFFECTIVENESS OF BILINGUAL EDUCATION

AIR	IBI
ENGLISH LANGUAGE ARTS	ENGLISH LANGUAGE ARTS
"In general, across grades, when total Title VII and non-Title VII comparisons were made, the Title VII students were performing in English worse than the non-Title VII students."[3]	On a vocabulary test in English, when Title VII children were compared to the comparison group of the same age and language dominance, Title VII children scored significantly higher at every age level. On a test of English Reading, when Title VII children were compared to the comparison group of the same age and language dominance, Title VII children scored significantly higher at every age level.
"Relative to national norms, Title VII Hispanic students, across grades, were performing at about the 20th percentile in English reading."[3]	Relative to national norms, Title VII Spanish dominant students were performing between the 37th and 47th percentile in English reading.
MATHEMATICS	MATHEMATICS
"In Mathematics, across grades, (Title VII children) were performing at about the same level as non-Title VII students."[3]	On a test of mathematics skills, when Title VII children were compared to the comparison group of the same age, children scored significantly higher at every age level.
"Relative to national norms, Title VII students were performing at about the 30th percentile in mathematics."[3]	Relative to national norms, Title VII students were performing between the 42nd and 68th percentile in mathematics.

The IBI program was not one of the Title VII programs included in the AIR study. It might be concluded, therefore, that IBI somehow represents one successful program in a generally unsuccessful educational effort. The author feels, however, that the real issues involved are: (1) what are the fundamental requirements for evaluation of effect in a bilingual program, and (2) are the conclusions of the two mentioned evaluations equally credible?

The U.S. Office of Education sets as its standard that the program must (1) be able to show reliable evidence of program effect, and (2) provide credible estimates of how the children would have performed without the program. To meet this later requirement, the AIR impact study used a comparison group model. Their original evaluation design called for school districts included in the study to nominate other classrooms in their district, or nearby districts, that would have children the same as those in the bilingual classrooms in all important variables. Their statement of this research design was as follows:

> Comparison classrooms nominated by each site were to contain students who would qualify for bilingual education and who were essentially the same ethnic background, linguistic competence [emphasis mine], and socioeconomic status as the students in Title VII project schools.[7]

The progress of these children over the year would provide the estimate of how children in the bilingual classrooms would have been expected to perform without the program. (Minor differences in pretest scores to be adjusted for statistically by analysis of covariance.)

The AIR study ran into problems with this evaluation design almost immediately. Out of the thirty-eight school districts included in the study, eighteen indicated they could not identify any comparison classrooms that met the criteria. The twenty remaining districts, however, did nominate classrooms that were used for the comparison group.

Teachers were asked to classify children as to their language dominance. At this point it became evident that the comparison classrooms had students of the same ethnic background as those in bilingual classrooms, but almost all of the students spoke only English (83%, or 1,349, out of the total sample in the Hispanic comparison group of 1,622 children).

Tables are provided of the classification by language dominance of Hispanic children who were present both fall 1975 and spring 1976 when the impact study testing was carried out. Based on teacher classification, 74% of the children in the bilingual classrooms were either non-English speaking or bilingual, compared to 17% in the comparison classrooms as seen in the following table:

Classification by Language Dominance of Hispanic Children Present Fall and Spring 1976[5]

	Title VII Classrooms		Non-Title VII Comparison Group	
Monolingual English	N=1,055	26%	N=1,349	83%
English Dominant Bilingual	N=1,908	47%	N= 188	12%
Spanish Dominant Bilingual	N= 273	7%	N= 8	* (less than 1%)
Spanish Monolingual	N= 860	21%	N= 77	5%

This would seem to present a serious dilemma, e.g., how to examine the effectiveness of an educational program for children who are bilingual or non-English speaking, when only the treatment group contains a significant number of bilingual or non-English speaking children.

The solution used in the AIR Impact Study was to change the evaluation design, eliminating the requirement that the two groups be compared on the basis of similar language competence. The final report issued in March 1978 is careful to leave language competence out of its descriptions of the variables on which a match was found:

> For each Title VII classroom selected for testing, district project personnel were asked to nominate non-Title VII classrooms (comparable students) within their own or within a nearby district whose students matched the Title VII student in terms of (a) ethnicity, (b) socioeconomic status, and (c) grade level [emphasis mine].[6]

Because the subsample of children who could be tested in Spanish was so small (under 5%) in the comparison classrooms, some tables of results in Spanish reading were included but no comparative analysis was attempted. The AIR report stated:

> It should be noted that the Impact Study's original analysis plans had specified the stratification of comparative analyses by judged language dominance group (i.e., separate analyses for groups of students who completed similar combinations of test and questionnaire instruments). However, this approach was not feasible in view of the small number of students in the non-Title VII comparison classrooms who were given test questionnaires developed for Spanish monolingual students or for Spanish dominant bilingual students.[7]

Spanish dominant children were given tests below their actual grade level for English reading. Using expanded scale scores, data were then pooled across all language dominance groups for the tests of program effect. This meant that the "credible estimate of how children would have performed without the bilingual program" was based on the gains made in "other" classrooms by children who were overwhelmingly English monolingual (85% in each of grades 2, 3, and 4). With this evidence the AIR Impact Study reports what they consider to be a "puzzling finding":

> In English Language Arts, participation in an ESEA Title VII Spanish/English bilingual education project did not appear to produce gains in student achievement over and above what would be expected had the students been assigned to a traditional classroom. In fact, in several grades, the non-Title VII students made slightly greater gains in English Language Arts. (A partial explanation of this puzzling finding may be substantially greater amount of time devoted to Spanish Language Arts instruction in the Title VII classrooms.)[8]

The IBI data are based on an evaluation design in which the evidence for how project students would have performed without the bilingual project is based on the accumulation of actual pretests of students who enrolled in the project at different ages, built up over a period of years. In this process a child who entered the program at age four would have a pretest which went into the "project norm group" data bank for four-year-olds. Subsequently, at age five this child's posttests would be compared to the project norm group for five-year-olds (made up

from pretests of other children who entered the program at age five). Because the project enrolls more new children at the younger ages the norm group at the upper age levels had to be expanded through some testing in a neighboring school district. The children in this testing met exactly the enrollment criteria used to enroll project children.

This type of comparison group provides an automatic match on such factors as ethnicity and socioeconomic level. Age, rather than grade level, is used as the basis for comparative analysis to overcome the problem of out-of-grade level placements which are very common for bilingual children.

The match on language competence between project posttest group and project norm group is accomplished by having all tests given in pairs--achievement tests and a language test together. The comparisons are then made separately by language dominance for curricular areas in which language dominance made a significant difference (reading and vocabulary in Spanish and English). It was found that for mathematics and handwriting, language dominance did not make a significant difference so the tests are pooled.

Attendance is kept individually in the IBI program and children are given posttests after each 100 day's attendance in the program. This means that children are being tested throughout the year. When an evaluation is done the tests are grouped by age and attendance period and primary language classification. Analysis is done to see if the superiority after differing periods of attendance is statistically significant when compared to the project norm group of the same

age, and the same language classification. The writer suggests that in answer to the question: "What are the fundamental requirements for evaluation of effect in a bilingual program?" that it would be reasonable to include the expectation that whatever comparisons are drawn be based on groups matched by language dominance.

The IBI data and the AIR data are on a totally different scale in terms of numbers of children tested and mass of data handled. AIR reports that "In all, over 11,500 students in 384 classrooms in 150 schools at 38 different sites across the United States were initially included in the Impact Study Sample." The IBI data are based on one program with three different sites. However, when numbers of children who are Spanish dominant are separately considered, the AIR study is based on 1,133 Spanish dominant children in the treatment group and 85 in the comparison group; the IBI data are based on over 600 tests of Spanish dominant children in the treatment group and over 800 in the project norm (comparison) group. In terms of providing information on children dominant in another language, the category of children who were the prime concern of Congress when it established the bilingual education program, there is much less difference in the scope of the two studies.

It is hoped that the concern for "answers" to the question of the effectiveness of bilingual education will begin to produce more programs with definitive data. The lack of such evaluation seems to this writer to be more related to the problems of how to find suitable tests, norm groups, and comparison groups for linguistic and ethnic minorities than it does to lack

of concern for evaluation. Some aspects of the evaluation design used by IBI may be helpful to other programs. The danger in the current mood of "panic" for evidence of effectiveness is that evaluations based on "not quite good enough" comparison groups, or "probably irrelevant" norms from tests standardized with English speaking children will produce a rash of conclusions that will distort rather than clarify the answer to the question: Does Bilingual Education Work?

REFERENCES AND NOTES

1. Cardenas, Jose A. Response I. In N. Epstein, Language, Ethnicity, and the Schools. Institute for Educational Leadership, the George Washington University, 1977.

2. Individualized Bilingual Instruction (IBI) was originally approved by the Joint Dissemination Review Panel in 1973 under the former project title: "Training Migrant Paraprofessionals in the Bilingual Mini Head Start."

3. American Institute for Research. Evaluation of the Impact of ESEA Title VII Spanish/English Bilingual Education Program. In National Dissemination and Assessment Center, Bilingual Education Paper Series, Vol. 2, No. 1, Aug. 1978, Los Angeles, California, p. 17.

4. American Institute for Research. Evaluation of the Impact of ESEA Title VII Spanish/English Bilingual Education Program, Volume I: Study Design and Interim Findings (AIR Report 48300-2/77-FR III), February 1977 (ERIC Report ED 138-090) II-2.

5. Ibid., Tables VI-B-17a, VI-B-18a, VI-B-19a, VI-B-20a, VI-B-21a, Appendix pages 123-127.

6. American Institutes for Research. Final Report: Evaluation of the Impact of ESEA Title VII Spanish/English Bilingual Education Program, March 1978, 5 pp.

7. American Institute for Research. Evaluation of the Impact of ESEA Title VII Spanish/English Bilingual Education Program, Volume I, page VI-17 (plate 180 of ERIC copy).

8. American Institute for Research, Final Report, p. 4.

APPENDIX

The following material was published in the annual evaluation of the IBI program for the 1977-78 program year. As it contains the data base from which the findings in this paper were taken, it is included here as an appendix.

GOAL 3. SPANISH SPEAKING CHILDREN WILL IMPROVE IN SPANISH.

THE NEED: The need for continued instruction in Spanish is based on the expressed wishes of the families whose children are served by the program. They want their children to keep their Spanish language capability and to improve it.

THE IBI CURRICULUM:

Curriculum Materials Used: IBI field tested the Spanish edition of DISTAR Language I, translated for IBI by Bilingual Mini Schools, a private corporation which has a license to sell the Spanish edition granted by SRA. The project selected these materials because they are very well programmed; that is, each new skill is practiced many times and then re-used as new skills appear. All skills are cumulative; they never disappear from the material. And because it is well programmed the children do not make frequent errors.

Spanish DISTAR is taught in a small group from a teacher presentation book which clearly outlines the dialogue used in teaching--a feature which is very helpful to a paraprofessional teacher. It involves a fast paced verbal exchange in which children would be expected to make from 150 to over 200 language responses during a 20-minute lesson period. Because the language responses require both phrases and whole sentences, the children learn to use the language with all its connecting words instead of learning an isolated vocabulary which often occurs with other language approaches.

The IBI program accepts very young children, age three and four, for whom a curriculum called "Pre-DISTAR" has been developed by the project. This curriculum begins with action responses and then requires one- or two-word responses. Children can use it who are not yet ready for whole sentence responses required in the regular DISTAR series. Pre-DISTAR is in both Spanish and English.

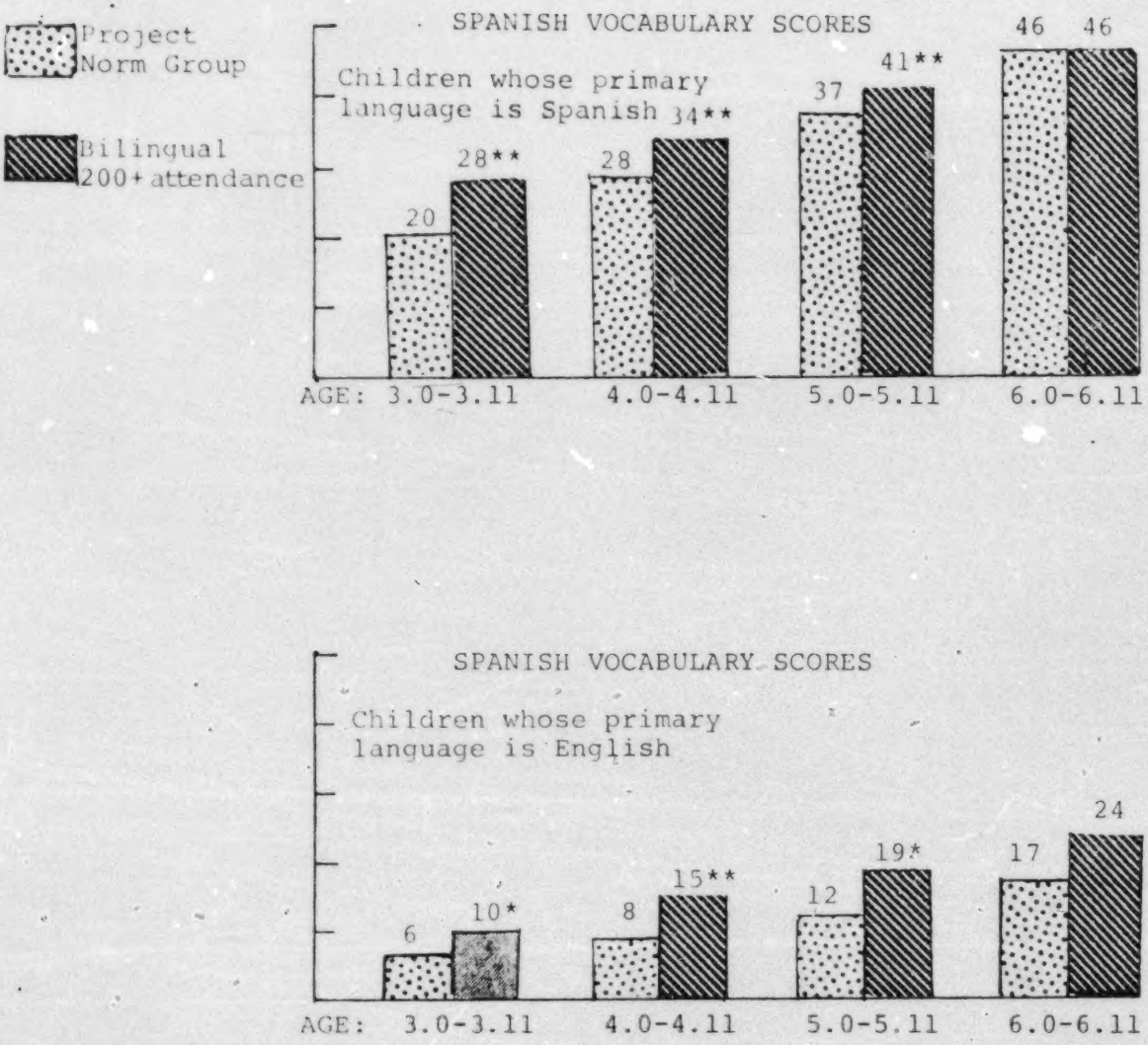
Lesson periods in oral Spanish of approximately 20 minutes per day are given all children in the preschool program and up to first grade. (DISTAR Level I is appropriate to the first

grade level. Level II is not yet available, so the project is waiting to extend the oral language program to the older school age children.)

In addition to the lessons in oral Spanish, all staff are bilingual and use both Spanish and English in the teaching of other subjects and in informal conversation throughout the day.

PROJECT RESULTS: Test results are reported separately by the child's primary language classification in Figure 3 which follows. These test results indicate that children whose primary language is Spanish have scores two to three times as great on this test than those whose primary language is English. However the average scores of children after 200 days attendance improve within each language group.

Table 6 in the technical appendix gives the statistical detail of the Spanish vocabulary test scores. As indicated, the superiority in Spanish of children after 200 days in the IBI program is statistically significant at age three, four and five, over the project norm group of the same age and language dominance. At age six the average score of English dominant children is superior to the norm group, but because of high variance within the group the difference is not statistically significant. The average score of six year old Spanish dominant children is the same for both project and norm group. One possible explanation of this is that only level one of the Language curriculum is available in Spanish (there are three levels in English and the curriculum is used up through third grade). The Spanish dominant children move through the curriculum in Spanish much more quickly than through the English curriculum and many who enter the IBI program as preschoolers have completed the available curriculum before age six. The project hopes to have the second and third levels of the curriculum available in the future which would enable the continuation of formal instruction in Spanish to a higher age level.



*The superiority of this score over the project norm group is statistically significant at the .01 level.

**Superiority over project norm group significant at .001 level.

Scores of children after 100 days attendance used because of an insufficient number of 200+ day tests for analysis at age three. Detailed test scores analysis is shown in Table 6 in the Technical Appendix.

FIGURE 3. MEAN RAW SCORES IN SPANISH VOCABULARY ON IBI TRANSLATION OF FORM B OF THE PEABODY PICTURE VOCABULARY TEST BY PERIOD OF ATTENDANCE, AND PRIMARY LANGUAGE CLASSIFICATION.

TO SUMMARIZE THE FINDINGS IN FIGURE 3:

1. BOTH SPANISH AND ENGLISH DOMINANT CHILDREN SHOWED GAINS IN SPANISH UNDER THE IBI CURRICULUM.

2. COMPARED TO THE AVERAGE SCORES OF THE PROJECT NORM GROUP OF THE SAME PRIMARY LANGUAGE CLASSIFICATION, THE SUPERIORITY AT AGE THREE, FOUR AND FIVE IS STATISTICALLY SIGNIFICANT.

GOAL 4. SPANISH SPEAKING CHILDREN WILL IMPROVE IN ENGLISH.

THE NEED: When children enter the IBI program they are given the Peabody Picture Vocabulary Test in English and in Spanish, using different forms of the test. Based on these entry tests for project children, 73% have a higher score in Spanish than in English. Of these children 46% have a score so low in English that it is within the range that could be obtained by guessing, and would be considered Spanish monolingual, as shown in Figure 4 (below). Even the children who are classified as English dominant are, in fact, very limited in their English skills.

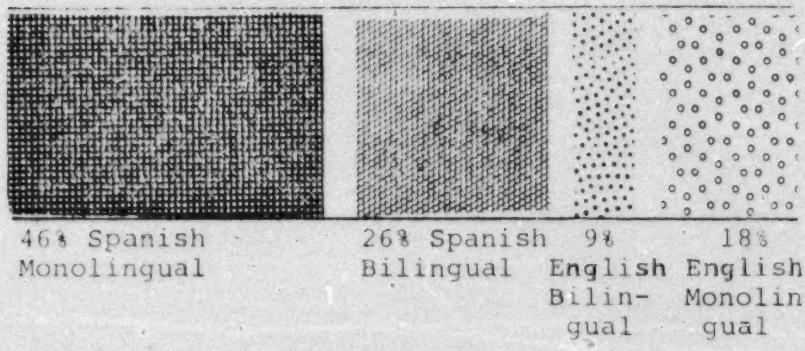


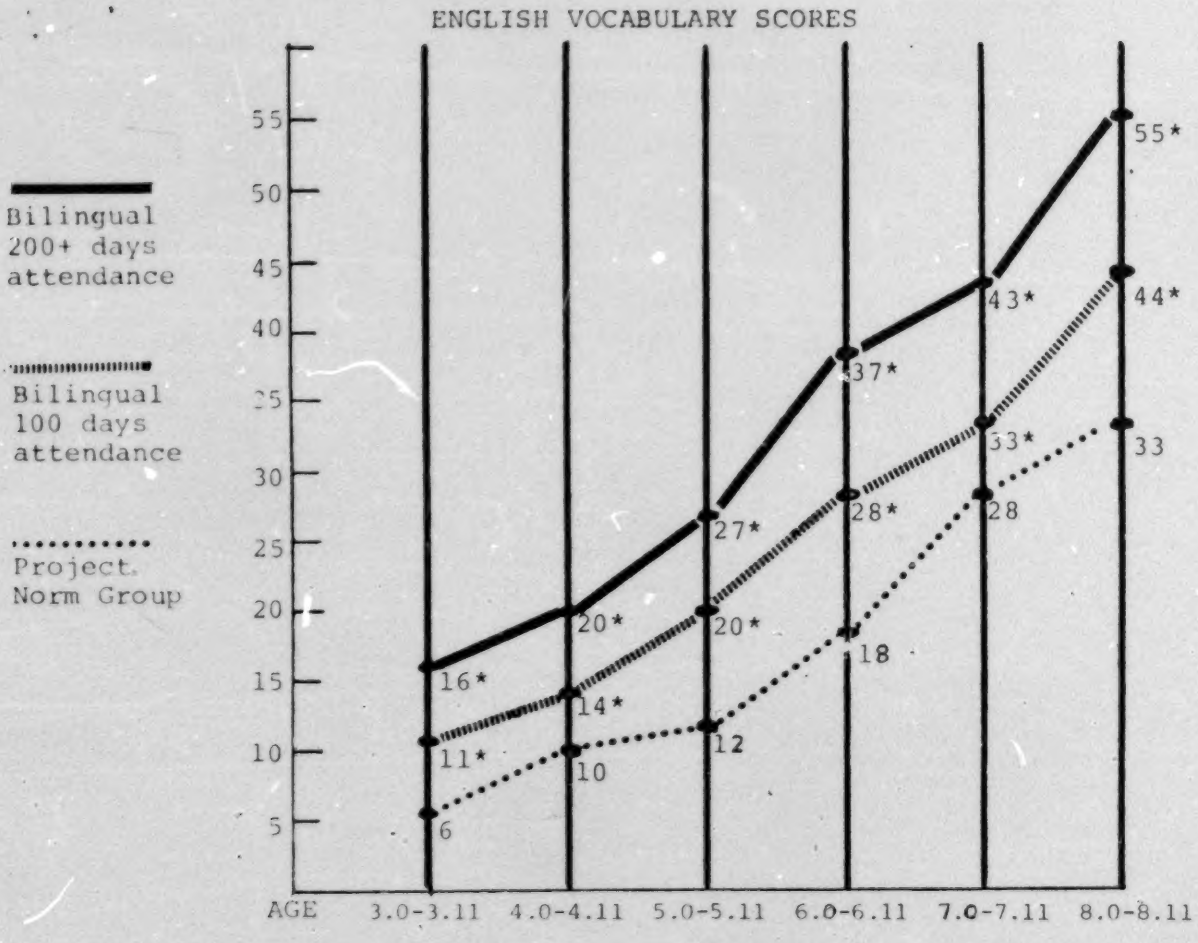
FIGURE 4. LANGUAGE CLASSIFICATION OF CHILDREN ENROLLED IN IBI PROGRAM THROUGH APRIL, 1978 BASED ON ENTERING SKILLS IN LANGUAGE.

IBI CURRICULUM: As described earlier both English and Spanish are taught each day (sometimes every other day depending on the length of time available with the children at a particular site). The DISTAR Language curriculum is used. This is carefully programmed, provides fast paced oral language practice, use of the language in context and in whole sentences. The context of the lessons involves language useful to understanding school usage vocabulary, reasoning skills, and readiness concepts.

The English Language curriculum is available from SRA at three levels and is used through third grade (the project only goes to third grade). Children are started on the program at the preschool ages. If the children are very young they started in the IBI

Pre-DISTAR curriculum. This shapes their ability to follow a teacher presentation in a small group, respond when asked, and starts with pointing responses leading into responses that require use of one or two words. This curriculum is in both Spanish and English. When children complete it they can usually move with no difficulty into the DISTAR language program.

In addition to the lessons in Spanish and English, both languages are used in the teaching of other academic subjects and informally interacting with children during child choice activity periods, lunch, and group activities such as singing, etc.



*The superiority of this score over that of the project norm group of the same age is statistically significant beyond the .001 level. Detailed test scores analysis is shown in Table 7 in the Technical Appendix.

FIGURE 5. ENGLISH VOCABULARY SCORES ON FORM A, PEABODY PICTURE VOCABULARY TEST, OF CHILDREN WHOSE PRIMARY LANGUAGE IS SPANISH, BY AGE AND ATTENDANCE GROUP.

TO SUMMARIZE THE FINDINGS IN FIGURE 5:

1. COMPARISON TO THE AVERAGE SCORES IN ENGLISH OF THE PROJECT NORM GROUP SHOWS SIGNIFICANT SUPERIORITY FOR BILINGUAL PROGRAM CHILDREN AT EVERY AGE LEVEL. IN ALL CASES THIS SUPERIORITY IS SUFFICIENT TO BE STATISTICALLY SIGNIFICANT BEYOND THE .001 LEVEL (E.G., AN EXCEEDINGLY SMALL POSSIBILITY THAT THIS MUCH DIFFERENCE WOULD OCCUR BY CHANCE).

2. CHILDREN AFTER 200 DAYS ATTENDANCE ARE MARKEDLY SUPERIOR TO THOSE TESTED AFTER ONLY 100 DAYS ATTENDANCE, INDICATING THAT THE GAINS IN ENGLISH ARE PROGRESSIVE THE LONGER THE PERIOD OF ATTENDANCE.

GOAL 5. CHILDREN WILL INCREASE THEIR SKILLS IN MATH.

THE NEED: IBI uses the Wide Range Achievement Test, math subtest, to measure children's skills in math. This is a nationally standardized test. Based on the national norm group, the average score of children before attending the bilingual program was at the 9th percentile at age five, and at the 19th percentile at ages six, seven and eight. This means that only 19 percent of the children in the national norm sample had scores this low or lower. It indicates that without the program this target population would have been expected to be far below national norms in math.

THE IBI CURRICULUM: The IBI project has developed and published a pre-math curriculum in Spanish and English for the pre-school age children. This is taught in small groups and in math related independent activities. Children learn numeral recognition, counting sequence, making sets, the concept of equal, etc.

After the pre-math series children begin using consumable workbooks published by Random House under the trade name of Singer "Sets and Numbers." This curriculum is based on set theory and is taught in units. IBI uses this series up through third grade (the upper level in this series uses hard cover books instead of the consumable workbooks).

These math materials can be completely individualized and a teacher can successfully work with a group all of whom may be on different pages or even in different books. Because of the disruptions in schooling experienced by a migrant population, the project used as one important criterion in the selection of curriculum materials whether the materials can be individualized.

The Sets and Numbers series are published in English only. This has presented no obstacle in their use by bilingual students taught by bilingual paraprofessional teachers. The teachers use both Spanish and English working with the students, for instructions, praise and correction. English terminology

is used for the mathematical terms and concepts. Since these are new concepts to the children in whichever language they are encountered, the English terminology readily becomes part of their vocabulary.

PROJECT RESULTS: The math subtest of the Wide Range Achievement Test was given to project children in English or in Spanish (project translation). It was found that language is not a major factor on the test and that there were no significant differences in pretest scores accumulated for the project norm group based on language dominance. Therefore the test scores in math were pooled for analysis.

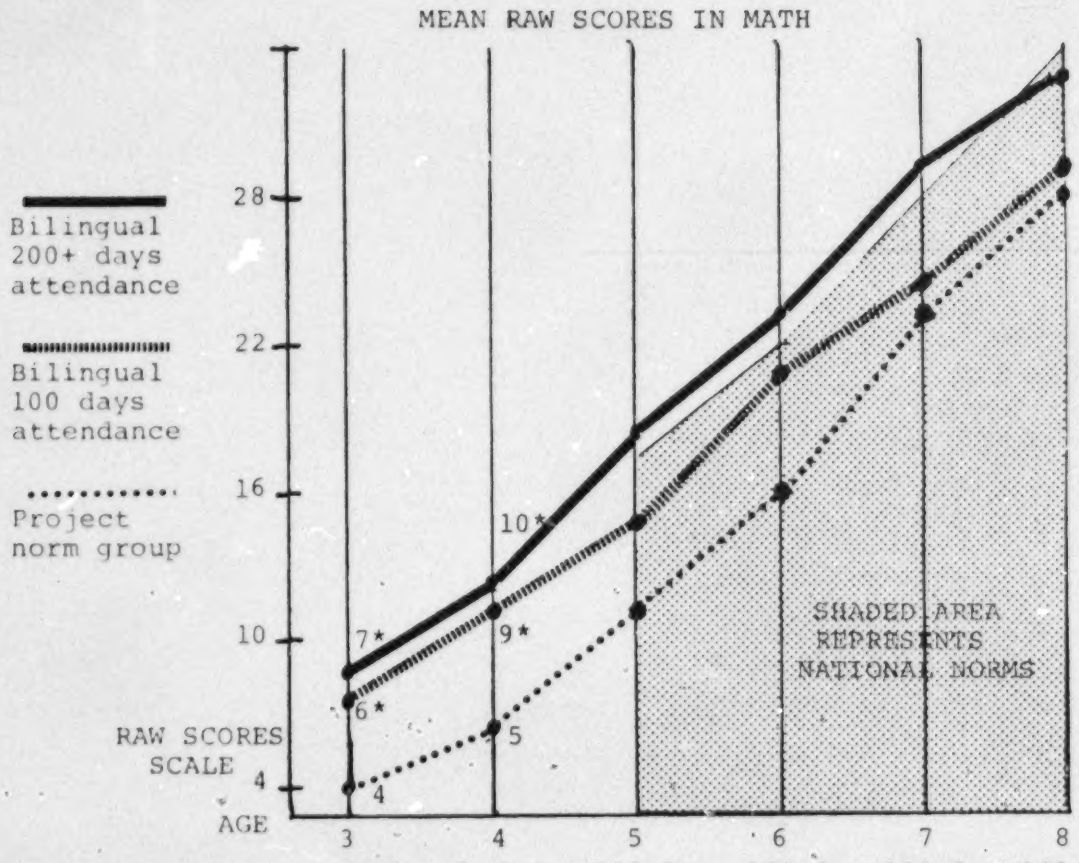
The graph on Figure 6 on the following page shows that children in the project norm group averaged scores far below the national norm. Short term project attendance (100 days) produced higher scores at every age level. The superiority after 100 days is statistically significant for children age three through six. The difference is not enough to be statistically significant at ages seven and eight. For statistical detail see Table 8 in the Technical Appendix.

Children who attended the IBI program for 200+ days had scores markedly higher than either the project norm group or children tested after 100 days attendance. The average score is also above national norms through age seven, slightly below the national norm at age eight. For statistical detail see Table 8 in the Technical Appendix.

Without the program the average score in math achieved by project children was below the 19th percentile by national norms. For students in the bilingual program the minimum 200 days, the percentile scores ranged from 42nd to the 68th percentile at different age levels.

The superiority of scores after 200+ days attendance was statistically significant at every age level, three through eight, beyond the .001 level of significance, (the possibility that this much superiority would occur by chance less than one in 1000). See Table 8, Technical Appendix.

Using the rule of thumb that a score more than half a standard deviation higher than the reference group would also be considered educationally significant, the superiority of children at every age level in the bilingual program with the minimum 200 days attendance would be considered educationally significant. At all but age eight the difference is more than one full standard deviation higher than the reference project norm group of the same age.



Mean Standard Scores	3	4	5	6	7	8
National Norm =	100.0		100.0	100.0	100.0	100.0
IBI 200 day =	104.6*		107.3*	101.3*	97.2*	
IBI 100 day =	94.7*		96.6*	91.8*	90.8	
Proj. Norm =	80.5		86.8	86.9	87.6	

*The superiority of this score over the project norm group of the same age is statistically significant beyond the .001 level. Detailed test scores analysis is shown in Table 8 in the Technical Appendix.

FIGURE 6. SCORES ON THE MATH SUBTEST OF THE WIDE RANGE ACHIEVEMENT TEST, BY AGE AND PERIOD OF ATTENDANCE IN IBI PROGRAM.

TO SUMMARIZE THE FINDINGS IN FIGURE 6:

1. WITHOUT THE BILINGUAL PROGRAM SCORES OF TARGET CHILDREN IN THE PROJECT NORM GROUP ARE FAR BELOW NATIONAL NORMS.
2. SHORT TERM ATTENDANCE (100 DAYS) IN THE BILINGUAL PROGRAM PRODUCES HIGHER AVERAGE SCORES THAN THE PROJECT NORM GROUP AT EVERY AGE LEVEL; HOWEVER STILL BELOW NATIONAL NORMS.
3. CHILDREN IN THE IBI PROGRAM FOR 200 OR MORE DAYS SCORE HIGHER AT EVERY GRADE LEVEL THAN EITHER THE PROJECT NORM GROUP, OR CHILDREN IN THE BILINGUAL PROGRAM ONLY 100 DAYS. THE SUPERIORITY OF THESE SCORES OVER THE PROJECT NORM GROUP AT EVERY AGE LEVEL IS STATISTICALLY SIGNIFICANT BEYOND THE .001 LEVEL.
4. THE IBI BILINGUAL GROUP WITH 200+ DAYS ATTENDANCE ALSO AVERAGES SCORES IN MATH ABOVE NATIONAL NORMS AT ALL AGE LEVELS EXCEPT AGE EIGHT.

GOAL 6. CHILDREN WILL GAIN READING SKILLS IN ENGLISH.

THE NEED: Children are tested on the reading subtest of the Wide Range Achievement Test for English reading. Tests are analyzed separately for children whose primary language is English and those whose primary language is Spanish because children differed significantly on their entry skills on this test based on language group.

As shown on Table 8 in the Technical Appendix, the entry level scores of children in the project norm group whose primary language was English were below the 25th percentile by national norms. The norm group scores of children whose primary language was Spanish all averaged below the 10th percentile by national norms (e.g., ten percent of children in the national norm sample had scores this low or lower.

THE IBI CURRICULUM: The preschool reading program in IBI is started at about age four using the University of Kansas Reading Primer, which was developed for the Behavior Analysis Head Start and Follow Through programs. The teacher holds a teacher presentation book and children sit in a semi-circle around the teacher. Children learn to respond to pictures going from left to right and from the top to the bottom of the page, and the basic skill of blending sounds. With these word attack skills the children are then able to begin in the Sullivan programmed reading series, published by McGraw-Hill. Once into the Sullivan program, the children can work at their own pace in an individualized program.

The Sullivan materials use a phonetic approach with a controlled vocabulary so that the beginning reader only has to remember one sound for each letter (except for a limited number of irregular words which are taught as sight words). This feature of the curriculum was found to be especially helpful to children whose primary language is Spanish. It means they do not have to cope with the many different sounds which letters take in English until after they have mastered the beginning reading skills of word attack, blending, etc.

The curriculum is programmed so the new material is introduced slowly with continuous review of what has been learned before. The picture illustrations are clear and uncluttered which is helpful to comprehension of a child for whom English is a second language.

The teacher circulates children working in a small group at a table. Workbooks are starred by the teacher for new sounds and children raise their hands for assistance when they come to a star. This enables the teacher to help the child with the new sound on a one to one basis. The child repeats the sound so the teacher can check that the child is both hearing and making the correct sound when it is first encountered. This avoids the confusion Spanish speaking children may have in distinguishing sounds in English that are not found in the Spanish language. The teachers are also trained to ask many comprehension questions as they check children's work, and can make explanations in Spanish if there is a comprehension problem.

A teacher is easily able to work with a small group of children each working on a different page, even in different books (there are 23 books in the Sullivan reading series).

PROJECT RESULTS: Figure 7 which follow illustrates the effect of the IBI program on the English reading scores of children whose primary language is Spanish. The shaded area represents the national norms for this test at each age level. Line A represents the scores of the project norm group, i.e., the expected score by target group children without benefit of the bilingual program. As can be seen in the graph the project norm group's score are far below the national norms, and the gap widens as the children get older.

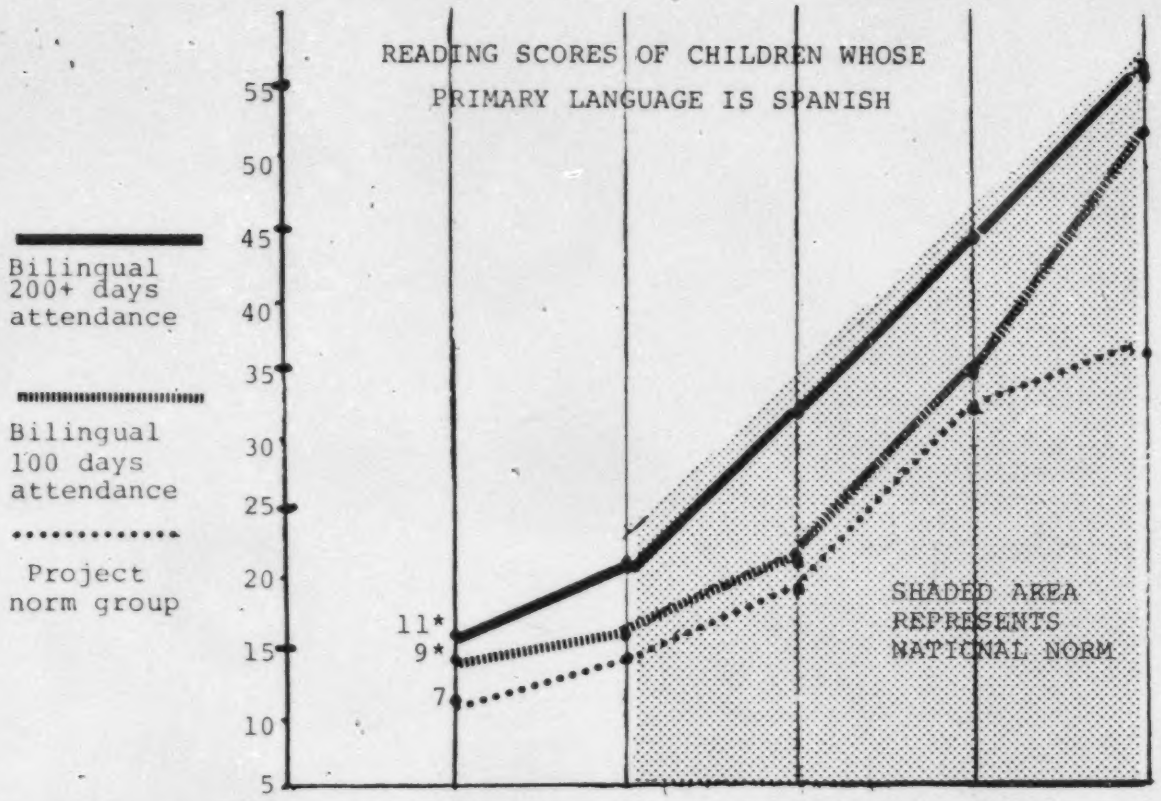
Children in the IBI program 100 days score higher at every age level than the project norm group. Children in the IBI program for 200 or more days have average scores higher than the 100 day test group, and much higher than the project norm group. The statistical detail for these scores is given in Table 9 in the Technical Appendix. The superiority of the

children in the bilingual program 200+ days over the project norm group is statistically significant beyond the .001 level in each age group.

At all age levels the score of children in the bilingual program is somewhat below national norms; however the average score comes closer to the national norm as the children get older. The mean standard scores for each group are reported under the graph in Figure 7. The national norm in standard scores would be 100 for each age level. As noted on the graph, the IBI children in the 200 day test group had a mean standard score of 95 at ages five and six, 97 at age seven, and 99.4 at age eight.

In terms of percentile scores the project norm group scores were at the 8th and 9th percentile in reference to the national norms for this test. The bilingual program children after 100 days averaged scores between the 14th and 19th percentile. After 200+ days, the scores of children ranged between the 37th percentile at age five and the 47th percentile at age eight.

Figure 8, which follows, and Table 9 in the Technical Appendix, present comparable data for children in the IBI program whose primary language is English. IBI accepts relatively few children who are English dominant and these almost entirely at the preschool or kindergarten level. For the ages and attendance groups in which there were a sufficient number of tests for analysis (a minimum of ten) the pattern is similar, only slightly higher than the corresponding scores for Spanish dominant children. The expected level of scores without the program, represented by the project norm group, is far below national norms; 19th percentile for age five and 25th percentile at age seven. Group mean scores are higher after 100 days in the IBI program; 39th percentile at age five and 30th percentile at age six. The IBI children with at least 200 days had the highest scores; 42nd percentile at age five and 55th percentile, e.g., above the national norms, by age six. The scores after 100 days were significantly higher (.05 level) than the project norm group. The superiority after 200 days was statistically significant beyond the .01 level.



Mean Standard Scores	National Norm:	100.0	100.0	100.0	100.0
	IBI 200+ day:	95.4*	95.2*	96.8*	99.4*
	IBI 100 day:	84.2	83.7	86.9	88.7
	Project Norm:	79.1	80.1	79.2	79.8

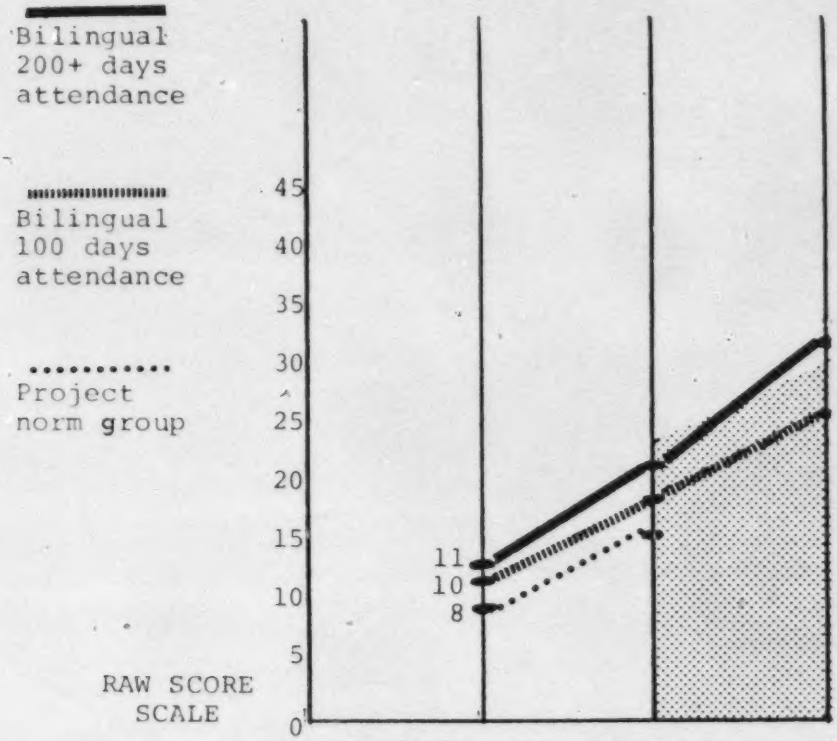
*The superiority of this score over the project norm group of the same age is statistically significant beyond the .001 level. Detailed test scores analysis is shown in Table 9 in the Technical Appendix.

FIGURE 7. MEAN RAW SCORES AND STANDARD SCORES OF IBI PROJECT CHILDREN BY AGE AND PERIOD OF ATTENDANCE ON THE READING SUBTEST OF THE WIDE RANGE ACHIEVEMENT TEST.

TO SUMMARIZE THE FINDINGS IN FIGURE 7:

1. PROJECT NORM GROUP SCORES INDICATE THAT WITHOUT THE BILINGUAL PROGRAM CHILDREN WHOSE PRIMARY LANGUAGE IS SPANISH WOULD BE EXPECTED TO HAVE READING SCORES IN ENGLISH FAR BELOW NATIONAL NORMS.
2. THE IBI CURRICULUM RESULTS IN HIGHER SCORES AT EVERY AGE LEVEL AFTER 100 DAYS, STILL HIGHER SCORES AFTER 200 DAYS, WITH THE SUPERIORITY AT 200 DAYS STATISTICALLY SIGNIFICANT BEYOND THE .001 LEVEL (POSSIBILITY OF CHANCE OCCURANCE LESS THAN 1 IN 1000).
3. AFTER 200 DAYS IN THE BILINGUAL PROGRAM, THE MEAN STANDARD SCORES OF CHILDREN ARE WITHIN THE 90-109 RANGE CONSIDERED AVERAGE BASED ON NATIONAL NORMS.

READING SCORES OF CHILDREN WHOSE PRIMARY LANGUAGE IS ENGLISH



Mean Standard Scores (Based on Norm of 100)	IBI 200 days:	97.6**	102.5
	IBI 100 days:	96.4*	92.3
	Proj. Norm:	87.3	(-)

**The superiority of this score over the project norm group of the same age is statistically significant at the .01 level. Detailed test scores analysis is shown in Table 9 in the Technical Appendix.

*Significant at .05 level.

(-) Less than 10 in group, too few for analysis.

FIGURE 8. RAW SCORES AND STANDARD SCORES OF IBI PROJECT CHILDREN BY AGE AND PERIODS OF ATTENDANCE ON THE READING SUBTEST OF THE WIDE RANGE ACHIEVEMENT TEST.

TO SUMMARIZE THE FINDINGS IN FIGURE 8:

1. THE PROJECT NORM GROUP SCORES FOR CHILDREN WHOSE PRIMARY LANGUAGE IS ENGLISH ARE IN THE LOW AVERAGE RANGE COMPARED TO THE NATIONAL NORMS OF THIS TEST.
2. AFTER 100 DAYS THE AVERAGE SCORES ARE HIGHER THAN THE NORM, HIGHER STILL AFTER 200 DAYS. THE SUPERIORITY AFTER EITHER 100 OR 200 DAYS IS STATISTICALLY SIGNIFICANT IN REFERENCE TO THE PROJECT NORM GROUP.
3. AFTER 200+ DAYS ATTENDANCE, CHILDREN AGE SIX HAVE A MEAN STANDARD SCORE OF 102.5 WHICH IS SLIGHTLY ABOVE THE NATIONAL MEAN SCORE OF 100.

4. As evident from Figure 8 and Figure 7 on the preceding page both children whose primary language is English and those whose primary language is Spanish appear to benefit from the IBI curriculum in improved reading skills.

TECHNICAL APPENDIX A

The main body of the final report was written for the general reader, and for clarity avoided technical detail. Since information on procedures for testing and data collection and explanatory footnotes for the analysis of test scores in the instructional component is of importance to a technical reader, that information has been reported here.

TESTING AND DATA COLLECTION AND ANALYSIS PROCEDURES

Training of Testing Personnel

Independent testers are used at each site--e.g., instructional or supervisory staff do not do testing. All testers are paraprofessional bilingual Mexican-Americans. Since more than 95% of the children served are Mexican-American, this testing staff has cultural identity for the overwhelming majority of children being tested.

Each tester is individually trained on each instrument and must demonstrate appropriate procedures in the testing of at least two children, under observation, as part of the training. Subsequent monitoring visits review testing procedures annually. Training and monitoring of testers is done by the evaluator and/or the resource trainer. Experienced testers are also used to assist in training new testers.

Testing Schedule

All children are pretested before attending the program for 30 days--in practice as soon as possible after their initial enrollment. This pretesting includes all of the instruments used in the evaluation appropriate to children of their age. It also includes the project achievement tests which are used to help determine initial placement of children in project materials. The IBI program serves children continuously over the 12 months of the year. Since new children can and do enroll in the program every month during the year, pretests are also given during every month throughout the year.

Repeat testing on standardized test instruments is done after an individual child has attended for a period of 100 days, i.e., the child is tested at 100, 200, 300, etc., days. Attendance data are kept individually for each child and at the end of each month a list is made of every child who that month passed a testing interval in his attendance. This list is forwarded to the tester at each site who then administers the appropriate tests for that child. Every month some children at each center reach a testing point in their program attendance, so posttests are also given during every

month throughout the year. Table 3 shows the statistics of the distribution of testing throughout the year for two of the project tests. The difference in total number of tests is because the PPVT has been in use by the project longer than the WRAT and more tests have been accumulated. These data indicate that the testing has been fairly evenly distributed throughout the year, and that the time of testing distribution is parallel for the test analysis groups utilized.

TABLE 3. Distribution Statistics on Month of Testing

	Peabody Picture Vocabulary Test				Wide Range Achievement Test			
Test Group	Number	Median	Mean	St.Dev.	Number	Median	Mean	St.Dev.
IBI Norm	887	5.80	5.94	3.89	679	5.89	5.95	3.90
200+ days	694	5.94	6.41	3.49	506	5.17	5.94	3.65

This testing schedule is more complicated than the usual evaluation procedure of doing mass testing at two calendar points. It was devised because of the unique requirements of evaluating a migrant population who have spotty attendance and who come and go at different times during the year. When the Texas Education Agency reports on child progress between tests administered in October and other tests administered in April of the school year it is not uncommon to have both pre- and posttests on less than 50% of the children enrolled in the program. Some children are there for pretests but leave before posttests are given. Other children come too late for pretests but are there at the posttesting dates. Some children enter late and leave early and aren't there for either testing point. This is a very common pattern in the evaluation of migrant programs--test information presented that represents a very unsatisfactory percentage of the project group it is meant to evaluate.

Another problem encountered in use of calendar date testing with a migrant population results from their uneven attendance patterns. Of two children tested on the same date in April, it would not be at all uncommon for one to have attended 50 days total prior to that testing point, another child to have attended 90 days. This puts the evaluator in the position of comparing child progress between tests which represent two very different amounts of exposure to the program being evaluated.

By testing based on individual attendance records, each of the posttests used in this evaluation represent a known amount of project intervention. At evaluation points the accumulated tests are then subgrouped by age and the period of attendance at the time of testing for purposes of analysis.

Selection of Tests

The IBI program selected the Peabody Picture Vocabulary Test, The Cooperative Preschool Inventory, and the Wide Range Achievement Test in part because these nationally standardized test instruments were among those most widely used by USOE in educational research and evaluation. The Cooperative Preschool Inventory was specifically developed for a measurement of the effects of Head Start programs and was utilized in many evaluative efforts measuring the impact of this nation-wide program of preschool education. The Peabody Picture Vocabulary Test similarly has been widely used in research on preschool programs. The Wide Range Achievement Test was one of two standardized achievement tests utilized in the massive Follow Through studies.

In order to use the tests, however, it was necessary to translate sections into Spanish. Tests are individually administered in the IBI program because of attendance cycle testing--this affects the math portion of the WRAT which was standardized under group administration. None of the tests was normed with a group even close to the IBI composition as to ethnicity or language use. For all of these reasons, the basic evaluation design has been kept internal to the project. The goals are all stated in terms of the project norm group, which is matched by ethnicity and language use to the project attendance groups. The language and circumstances for test administration are held constant in the testing of these two groups--indeed the testers are the same. The references, where made, to national norms are added so that there is some external interpretive power to the data. But all of these reservations about the application of the national norms restricts the interpretation of the data. However, this would be true of any bilingual program.

Analysis Procedures

The project norm group is used as a measure of the probable achievement level of project children without benefit of the program. When a child enrolls in the program he is pretested. If the pretest was given before the child had attended the program for as long as 30 days, it is put into the project norm group for children at that age level. The project has an enrollment policy which permits children to start the program at different ages, e.g., some start when they are three, others when they are four, others at five, or six. By accumulating pretests the project has been able to develop its norm group for all ages on all tests. The size of the norm group is increasing constantly as new pretests are accumulated.

In the past all evaluations have used a norm group obtained as described in the previous paragraph. As of 1977, the norm group for the Peabody Picture Vocabulary Test and for Wide Range Achievement Test was expanded to include the tests of children from a neighboring town to La Grulla, Texas. These children had been tested for the previous three years to provide a comparison group of mobile migrant children to compare to the children in just our

mobile component using the two tests mentioned, the PPVT and WRAT. The reason for testing a comparison group instead of just using the project norm group for this special study (see instructional goal nine) is that the project norm group includes pretest scores on some children from permanent sites who do not migrate. In the special study we wanted to hold the factor of migration constant: only project children who migrated were tested, only comparison group children who migrated were tested.

The community where comparison group tests were given and La Grulla have about the same socioeconomic level--most families earn their yearly income from the migration period doing seasonal farm work. Both communities are Spanish dominant, located on the Rio Grande River where television stations beam in Spanish language programs, many radio stations are in Spanish, the usual language of casual conversation in town or school would be Spanish. As a further check on the comparability of the two population groups, statistical analysis was done for two years in a row comparing the mean test scores by grade level of the comparison group and the pretest scores of children at the same level in the project norm group. No significant differences were found. This was interpreted to mean that the children from the neighboring town were, in fact, the same population group as project children and their scores are like those of the children pretested for this program.

Most pretest scores for children in this project were at lower age levels (three, four, five) since the project attempts to get children started as young as possible. Therefore, the addition of scores from the comparison group to the project norm group enlarged the size of the norm group at the school age level where is has been the smallest. Having a larger norm group at these upper ages allows greater stability for statistical analysis.

The project evaluation design calls for comparison of the mean scores of project children posttested after 100 or 200 days to children in the project norm group of the same age. A t-test of statistical significance is run between the project posttest attendance groups and the project norm group of the same age. Statistical analysis is only done when subgroups to be compared have a minimum size of ten.

For this evaluation all tests data were put onto computer cards and the t-test analysis was done using the SPSS computer package. The readout using this computer analysis gives the t-value using separate or pooled variances. If a significant difference existed in the two variances the separate variance t was reported; if the two variances were not significantly different the pooled variance t was reported in the tables of detailed project findings which follows in this technical appendix.

Most objectives are based on comparison of the children with over 200 days attendance and the project norm group. Because it is difficult for a child to accumulate 200 days of attendance all at the same age level, there is almost complete independence of the two groups used in the analysis. Children with 200 days at age four will more than likely have started the program at age

three, for example. They will be compared to the pretest scores of other project children who started the program at age four.

A very small percentage of the children have both pretest and posttest scores at the same age level. Analysis was therefore done including and excluding this small overlap group. The conclusion was that any bias created by this lack of complete independence of groups acted against the project in reducing the likelihood of a finding of significance.

The test analysis tables report long term program effect through the cumulative analysis, e.g., all tests through April 1978 which was used as the cut-off date for this evaluation. The current year program data would be posttests given in the period since the last analysis, e.g., May, 1977 through April, 1978. The cumulative norm group is used for tests of significance of difference between means for both long term cumulative and current year analysis groups. The t-value reported in each table is between the posttest attendance group, and the project norm group of the same age category.

To maintain a high level of accuracy in handling of test data the following precautions are taken. Each tester scores tests she administers. All tests are checked and rescored upon receipt. After preparation of the computer cards a readout is obtained and two persons recheck the accuracy before the decks are used. After analysis has been done, the N's of subgroups are rechecked against the project data entry records.

The project has employed independent evaluation specialists who have reexamined all analysis procedures and validated claims. The evaluation division of Northwest Educational Research Laboratories in Portland performed this evaluation review and audit through 1974 and Technical Assistance Services of Seattle through 1978.

1.3 SPANISH VOCABULARY

TABLE 6. Peabody Picture Vocabulary Test Scores, Form B in Spanish

Children Whose Primary Language is Spanish

Age and Attendance	Number	Mean	St. Dev.	t-value*	2-tail prob.
AGE 3.0-3.11					
IBI norm	221	19.950	8.337		
100 days	81	26.000	8.868	5.52	0.000
200+ days	15	28.267	6.984	4.39	0.001
AGE 4.0-4.11					
IBI norm	132	28.402	8.954		
100 days	145	31.214	9.023	2.60	0.10
200+ days	134	34.090	8.075	5.44	0.00
AGE 5.0-5.11					
IBI norm	98	36.735	9.390		
100 days	76	40.632	9.005	2.76	0.006
200+ days	176	40.910	7.781	3.75	0.000
AGE 6.0-6.11					
IBI norm	66	46.742	10.164		
100 days	28	43.393	9.005	-1.47	0.145
200+ days	108	46.639	8.111	-0.07	0.944

Children Whose Primary Language is English

AGE 3.0-3.11					
IBI norm	54	6.091	6.331		
100 days	30	10.000	6.486	2.69	0.009
200+ days	5	(too few for analysis)			
AGE 4.0-4.11					
IBI norm	61	8.148	10.076		
100 days	43	14.279	11.724	2.86	0.005
200+ days	44	14.659	9.829	3.30	0.001
AGE 5.0-5.11					
IBI norm	26	11.885	11.687		
100 days	33	14.970	14.668	0.88	0.385
200+ days	62	18.984	12.517	2.47	0.015
AGE 6.0-6.11					
IBI norm	13	17.077	17.380		
100 days	11	25.000	18.836	1.07	0.296
200+ days	27	24.482	16.568	1.30	0.200

*The t-value compares the means of the posttest attendance groups to the project norm group of the same age and language classification.

1.4 ENGLISH VOCABULARY

TABLE 7. PEABODY PICTURE VOCABULARY TEST SCORES IN ENGLISH, FORM A

Cumulative Analysis

Test Group: Children Whose Primary Language is Spanish

Age and Attendance	Number	Mean	St. Dev.	t-value	2-tail prob.
AGE 3.0-3.11					
IBI norm	221	6.448	4.211		
100 days	82	10.866	6.478	5.74	0.000
200+ days	15	16.267	8.598	4.39	0.001
AGE 4.0-4.11					
IBI norm	132	9.879	7.185		
100 days	145	14.255	8.075	4.75	0.000
200+ days	134	20.425	12.422	8.49	0.000
AGE 5.0-5.11					
IBI norm	98	12.398	9.585		
100 days	76	20.145	13.637	4.21	0.000
200+ days	178	27.292	15.220	9.95	0.000
AGE 6.0-6.11					
IBI norm	66	17.682	11.911		
100 days	28	28.000	14.684	3.58	0.001
200+ days	108	37.583	15.351	9.56	0.000
AGE 7.0-7.11					
IBI norm	67	27.896	15.818		
100 days	6	33.333	20.520	(too few for analysis)	
200+ days	52	42.885	12.502	5.61	0.000
AGE 8.0-8.11					
IBI norm	46	33.261	16.686		
100 days	3	44.333	11.590	(too few for analysis)	
200+ days	37	55.081	9.722	7.44	0.000

Current Year Analysis*

AGE 3.0-3.11					
100 days	10	13.700	8.908	2.56	0.031
AGE 4.0-4.11					
100 days	35	12.657	5.841	2.11	0.036
200+ days	28	21.500	14.688	4.08	0.000
AGE 5.0-5.11					
100 days	13	23.615	14.385	2.73	0.017
200+ days	42	26.310	15.371	5.43	0.000

TABLE 7. (continued)

AGE 6.0-6.11					
200+ days	15	40.447	14.845	6.52	0.000
AGE 7.0-7.11					
200+ days	20	43.400	13.100	4.73	0.000
AGE 8.0-8.11					
200+ days	15	58.467	10.246	6.98	0.000

*Too few for current year analysis in 200+ days at age 3, and 100 day groups at ages 6, 7, and 8.

TABLE 8. WIDE RANGE ACHIEVEMENT TEST, MATH SUBTEST SCORES

Cumulative Analysis

Age and Attendance	Number	Mean St. Score	Nat'l %ile	Raw Sc. Mean	St. Dev.	t-value	2-tail prob.
AGE 3							
IBI norm	183	--	--	2.929	2.046		
100 days	74	--	--	5.851	2.713	8.35	0.000
200+ days	12	--	--	7.250	3.361	4.40	0.001
AGE 4							
IBI norm	134	--	--	4.754	2.895		
100 days	155	--	--	8.439	3.754	9.41	0.000
200+ days	128	--	--	10.422	3.456	14.36	0.000
AGE 5							
IBI norm	102	80.461	9th	8.804	4.115		
100 days	95	94.674	37th	12.295	4.368	5.78	0.000
200+ days	176	104.591	63rd	14.972	3.779	12.69	0.000
AGE 6							
IBI norm	74	86.811	19th	13.446	4.852		
100 days	44	96.591	42nd	16.500	5.028	3.26	0.001
200+ days	92	107.348	68th	19.250	4.510	7.97	0.000
AGE 7							
IBI norm	79	86.924	19th	19.127	4.876		
100 days	24	20.167	30th	20.167	5.079	0.91	0.367
200+ days	51	101.275	53rd	24.412	3.093	7.56	0.000
AGE 8							
IBI norm	52	87.615	19th	22.981	4.676		
100 days	10	90.800	27th	24.300	2.263	1.37	0.183
200+ days	33	97.152	42nd	27.242	1.921	5.84	0.000

Current Year Analysis

AGE 3							
100 days	12	--	--	5.751	2.221	4.60	0.000
200+ days	4	(too few for analysis)					
AGE 4							
100 days	39	--	--	7.103	3.299	4.32	0.000
200+ days	33	--	--	10.394	3.316	9.73	0.000
AGE 5							
100 days	17	94.824	37th	12.235	3.898	3.21	0.002
200+ days	57	101.597	53rd	14.211	3.406	8.43	0.000
AGE 6							
100 days	10	82.400	12th	11.800	5.996	-0.98	0.330
200+ days	27	103.852	58th	18.074	4.178	4.39	0.000

TABLE 8. (continued)

AGE 7							
100 days	5	(too few for analysis)					
200+ days	18	102.833	58th	25.000	1.910	5.01	0.000
AGE 8							
100 days	0						
200+ days	18	97.889	45th	27.556	2.064	5.64	0.000

TABLE 9. WIDE RANGE ACHIEVEMENT TEST SCORES, READING SUBTEST

Cumulative Analysis

CHILDREN WHOSE PRIMARY LANGUAGE IS SPANISH

Age and Attendance	Number	Stand. Score	Nat'l %ile	Mean Raw Score	St. Dev.	t-value	2-tail prob.
AGE 4							
IBI norm	84	--	--	6.786	4.084		
100 days	123	--	--	8.626	3.463	3.49	0.001
200+ days	94	--	--	10.564	3.840	6.36	0.000
AGE 5							
IBI norm	75	79.080	8th	9.200	3.792		
100 days	71	84.225	14th	11.113	4.982	2.60	0.010
200+ days	126	95.437	37th	16.405	7.096	9.37	0.000
AGE 6							
IBI norm	62	80.145	9th	14.258	6.763		
100 days	33	83.667	14th	16.242	7.429	1.32	0.192
200+ days	77	95.221	37th	23.610	10.219	6.46	0.000
AGE 7							
IBI norm	66	79.212	8th	23.439	10.043		
100 days	15	86.867	19th	27.933	15.434	1.08	0.297
200+ days	43	96.767	42nd	38.884	14.394	6.13	0.000
AGE 8							
IBI norm	45	79.822	9th	31.111	11.871		
100 days	6	(too few for analysis)					
200+ days	30	99.400	47th	50.733	10.859	7.25	0.000

CHILDREN WHOSE PRIMARY LANGUAGE IS ENGLISH

AGE 4							
IBI norm	34	--	--	7.765	4.046		
100 days	32	--	--	9.563	3.141	2.01	0.049
200+ days	34	--	--	11.382	7.765	3.26	0.002
AGE 5							
IBI norm	19	87.316	19th	12.579	6.354		
100 days	24	96.375	39th	16.250	5.795	1.98	0.055
200+ days	49	97.633	42nd	17.571	6.110	2.99	0.004
AGE 6							
IBI norm	8	(too few for analysis)					
100 days	11	92.273	30th	22.455	8.858	(no norm group)	
200+ days	15	102.467	55th	27.467	8.927	(no norm group)	
AGE 7							
IBI norm	11	90.545	27th	35.091	13.389		
100 days	9	(too few for analysis)					
200+ days	8	(too few for analysis)					

TABLE 9. (continued)

Current Year Analysis

PRIMARY LANGUAGE SPANISH							
AGE 4 (200+)	26	--	--	9.539	2.996	3.18	0.002
AGE 5 (200+)	44	92.500	32nd	14.500	5.337	5.79	0.000
AGE 6 (200+)	23	94.304	34th	22.870	9.593	3.96	0.000
AGE 7 (200+)	14	97.643	45th	39.286	14.398	3.92	0.001
AGE 8 (200+)	18	103.278	58th	54.111	10.649	7.14	0.000
PRIMARY LANGUAGE ENGLISH							
AGE 5 (200+)	13	96.231	39th	16.846	6.388	1.86	0.072

Too few for analysis at any other ages.